PARADOXES OF EARLY SET THEORY

Cristina Criste

"Babes-Bolyai" University,
Facultaty of History and Filosophy Cluj-Napoca

Abstract: *The paradoxes of set theory, usually connected with the names of Russell, Zermelo, or Burali-Forti, are concerned with the contradictory notion of 'the set of all sets'. A further consideration of the topic, going back to the mathematical and even philosophical investigations of Cantor himself, could prompt interesting historical reassessments and reveal new insights into the configuration of set theory. I will show that Cantor's adherence to some concepts such as absolute actual infinite or consistent and inconsistent multiplicities, together with the rejection of the principle of comprehension are fundamentally connected with his outlook on the paradoxes and also with the way he introduced and configured the basic concepts of set theory – the primitive notion of set, the transfinite numbers, the future principle of well-ordering – and their role as the foundations of mathematics.*

Key words: *paradox, set theory, Cantor, Russell, infinite.*

INTRODUCTION

From the early stages of set theory, the paradox of 'the set of all sets' has known a multitude of versions, the most famous of all being that associated with the name of Bertrand Russell. The Russell's paradox is a logical and set theoretical paradox determined by considering the set of all sets that are not members of themselves: a set would be a member of itself if and only if it is not a member of itself.

But the history and development of Russell's paradox are not completely identical with the problems encountered by Cantor in the constitution and modulation of the notion of set and its further integration in the Zermelo-Fraenkel-Skolem axiomatization of set theory.

In his *Introduction to mathematical philosophy*, Russell admits that when he "first came upon" the contradiction, in 1901, he was actually attempting "to

An. Inst. de Ist. „G. Barițiu" din Cluj-Napoca, Series Humanistica, tom XIV, 2016, p. 241-259

discover some flaw in Cantor's proof that there is no greatest cardinal" (Russell, 1919/2010, p. 136).

That being said, some further historical nuances are to be taken into consideration. Russell 'discovered' the antinomy sometimes in 1901 (later accounts do not all coincide), while working on the *Principles of Mathematics* (published in 1903). Acknowledging its baneful effect on Frege's attempt to provide a theory of the foundations of mathematics, he wrote to the latter in 1902, emphasizing the inconsistencies of the axioms used to formalize Frege's logic. The paradox was later published in *The Principles of Mathematics* (1903).

> I may mention that I was led to it in the endeavor to reconcile Cantor's proof that there can be no greatest cardinal number with the very plausible supposition that the class of all terms (which we have seen to be essential to all formal propositions) has necessarily the greatest possible number of members. (Russell, 1903/1996, p. 101)

Regarding Frege's investigations, the contradiction is determined by the existence of what Frege called the "basic law V" (Frege, 1893/1964), the first explicit formulation of the principle of comprehension, introduced as an axiom and considered a logical principle, from which (together with the other "basic laws") to derive fundamental axioms and theorems of number theory. It is based on the notion of the course-of-values (*Werthverlauf*) of a function, with the extension of a concept as a special case. It must be pointed out that, for Frege, the function is not an object, but something incomplete, unsaturated and, as a consequence, it requires an object as an argument. According to *Rule V*, two functions f and g determine the same course-of-values if and only if, for the same argument x, they have the same value, i.e. $f(x) = g(x)$.

The contradiction implied by *Rule V* is that the function would have itself as argument: since the predicate is a function $\varphi(\xi)$ and if we take the function to be a concept with extension (or the corresponding class, *Werthverlauf*) "$\dot{\varepsilon}\varphi(\varepsilon)$", we would have a case in which the concept is predicated of its own extension. The function taken as a concept represents a special case of *Rule V*: two concepts determine the same course-of-value if and only if they have the same objects falling under each of them.

Frege admits that the contradiction affects the generalization of the transformation of an equality into an equality of the courses-of-values (which therefore "is not always permitted"), that *Rule V* is false and, because of that, we reach the loss of "not only the foundations of my arithmetic, but also the sole possible foundations of arithmetic" (Frege's letter to Russell, as cited in Heijenoort, 1967, pp. 127-128).

Russell's paradox is fundamentally connected with other results, some antedating Russell's results. Firstly, he acknowledges the contradiction "discovered" by Burali-Forti in 1897 that there isn't a greatest ordinal, while correctly rejecting some of the latter's inferences (Russell, 1903/1996, p. 327). The Burali-Forti paradox posits that if the set of all ordinals is well-ordered, it has an ordinal, but that ordinal would concomitantly be an element of the set of ordinals and greater than any ordinal in the set.

The article of Burali-Forti from 1897[1], to which, wrongfully, the "discovery" of the paradoxes was related to, did not even claim to have found a paradox of set theory. It only contained a proof by *reductio ad absurdum* of the fact that "there actually exist *transfinite numbers* (or *order types*) *a* and *b* such that *a* is not equal to *b*, not smaller than *b*, and not larger than *b*" (Burali-Forti, 1897, as cited in Heijenoort, 1967, p. 105), in other words, that the natural ordering of ordinals is just a partial ordering. Later that year, Cantor would demonstrate that the set of all ordinals is linearly ordered. The problem with Burali-Forti's article is that it is based on an erroneous reading of Cantor's notion of well-ordered sets[2], one that himself acknowledged in a note published eight months later[3]. Burali-Forti's article involves a fundamental concept in set theory and the difficulties around it: the well-ordering principle.

THE WELL-ORDERING PRINCIPLE

A or *the* fundamental question of set theory is whether each set can be well-ordered[4]. A well-ordered set is a special kind of simply-ordered set. The last one, now called totally-ordered set, represents an order relation on a set such that for any two different elements *a* and *b*, either $a < b$, or $b < a$. To be considered well-

[1] February 1897: "A Question on Transfinite Numbers", communicated at the 28 March 1897 meeting of the Circolo matematico di Palermo and published in its Rendiconti (Heijenoort, 1967, pp. 105-111).

[2] He introduces his own notion of "perfectly ordered sets", which is not identical with that of well-ordered sets: as he will point out, "every well-ordered class is also perfectly ordered, but not conversely" (Burali-Forti, 1897/1967a, 1897/1967b, p. 112).

[3] Under the title "On well-ordered classes", October 1897 (Heijenoort, 1967, pp. 111-112).

[4] Although equivalent, Cantor's characterization of the well-ordering principle is not quite the same as the now standard one. For Cantor, M is well-ordered by a relation $<$ if:

(i) M is linearly ordered by $<$;

(ii) M has a $<$-first element, m0; הּ

(iii) Whenever $N \subset M$ and $\exists\, m \in M - N\ \forall\, n \in N\ [n < m]$ then there is a $<$-smallest $m \in M - N$ such that $\forall n \in N\ [n<m]$. (G.168). Today, M is well-ordered by a relation $<$ if (i) it is totally ordered (that is, given any two elements, one is smaller), (ii) there is no infinite decreasing sequence, (iii) although there may be infinite increasing sequences, (iv) every non-empty subset of M has a least element.

ordered, an additional aspect must be fulfilled: every subset of the set must have a least element. But Cantor's characterization of well-ordering is not quite the same as the now standard one, as Hallett (1984, p.51), too, points out.

Cantor introduced the principle or hypothesis of well-ordering in 1883 as "a law of thought" (Cantor, 1883/2005[5], p. 886), without further explicit developments of its significance[6]. Although vague, this notion must be related with Cantor's view on the concept of set.

This *law of thought* assures the *determinate* relations among the elements of a well-defined set. In other words, when "well-defined", a set is given together with a linear ordering. In the new context of set theory, the distinction between the potential and the actual infinite corresponded to the distinction between the undetermined and the determined. In the case of an infinite set, the process of determination, by the laws of thought, acts together with the process of counting; they codetermine the ordinal, initially called *Anzahl* (Cantor, 1883/2005, p. 884).

Because of the definition of set, and the focus on their intellectually imposed unity, the finite and the infinite fall under the same 'construction' principles. Additionally, there is an epistemological neutrality in conceiving infinite or finite objects, according to the principles adopted by Cantor in the *Grundlagen*[7]. In those circumstances, *it is as difficult to understand the infinite as it is to consistently conceive the finite.*

A well-defined set already implies unity and therefore existence. Because it already has a linear order, it could be said that a well-defined set is intrinsically determined, with regard to its elements: (i) they exist in (ii) relations. And the idea of relation is here fundamental (also see Cantor, 1882/1932, p. 150).

The *law of thought* assures the immediate passage from the well-ordered set to the ordinal or *Anzahl*, but we do not know more than the fact that there is an image in the inner intuition, an ideal copy of the set, which offer both of them (and also to the power of the set) equal ontological value. The situation will be complicated by the theory of abstraction.

[5] For the rest of the text, Cantor's articles will be cited in the same way: the first year represents the first publication of the article (in German or French), while the second, the year of publication of an English translation, or Zermelo's edition of almost all of Cantor's work (Zermelo: *Georg Cantor – Gesammelte Abhandlungen mathematischen und philosophischen Inhalts,* 1932), when English translations are not available, or considered unsatisfactory. The page will correspond to the second edition quoted. The names of the articles can be found, in the chronological order of appearance, in the general bibliography at the end of the text.

[6] In a logicist tradition, a law of thought might be considered a logical principle.

[7] For example, "Omnia seu finita seu infinita definita sunt et except o Deo ab intellectu determinari possunt" [All things finite or infinite are definite and, God excepted, can be determined by the intellect.] (Cantor, 1883/2005, p. 891).

The order-types represent a developing of the notion of well-ordering, such that this principle of well-ordering conceived of as a law of thought makes possible the development of number-classes and the possibility to constitute a scale for comparing transfinite numbers. Given an ordinal α, each infinite cardinal also represents the power of the α^{th} number-class. In the aleph notation, every infinite cardinal is an aleph (the Aleph Theorem).

In view of the results from the *Grundlagen*, Cantor sends to Kronecker (in a letter from August 1884) a new definition of the transfinite numbers, one assuming well-ordered sets. The "law of thought" to which Cantor was referring presumed that the growing powers of the number classes are all transfinite cardinalities, guaranteeing the comparability of the cardinals (see Cantor 1883/2005, pp. 916-7, n. 2).

Eventually, Cantor realized that both well-ordering and the cardinal comparability demanded proofs. That being said, they remained open problems, and the supposed overviews of set theory provided by the *Beiträge* (Cantor, 1895/1897/2010) did not overtly emphasized the significance of well-ordering, only mentioning the cardinal comparability.

The notion of well-ordering is related to other fundamental set-theoretical principles or theorems. I will only emphasize here its association with the extensionality of sets, assumed by Cantor from the beginning: he uses m, m', m'',... for the elements of an *arbitrary* set M. Ignoring the unity of a set, we have distinct, discrete elements (Cantor, 1887-1888/1932, p. 380). *We subsequently choose and impose an order*. Is not entirely clear what is successive, since in the case of infinite sets, every choice implies *simultaneously* an infinity of elements.

But that doesn't mean that an infinite set is composed of parts. The 'reality' of the elements of sets is distinct from the 'reality' of the set that contains the given elements. Such an approach is similar to some of Spinoza's distinctions, a philosopher Cantor assiduously studied himself: since the essence of the Modes does not necessarily include existence, their "delimitation", conceived in terms of divisibility and measure (greater or less) is arbitrary. To attempt to do so with Substance is incompatible with its essence, hence the paradox in considering the extended Substance as made up of parts: the substance would be finite, annulling its very essence. When Duration, for example, is regarded abstractedly, namely, confused with Time, it is conceived as composed of moments and one gets to the paradox: "he can never understand how an hour, for instance, can pass by" (Letter 12, in Spinoza, 2002, p. 789).

And what is more, the well-ordering principle is basically connected to the attempts to offer a solution to the continuum hypothesis: cardinals must be arranged in a scale, therefore they must be compared. In 1915, Hartogs proved that the comparability of cardinal numbers representing sets depends and is equivalent to the well-ordering theorem.

All in all, one can agree with A. Kanamori (2009, p. 395) that “Zermelo's axioms can be construed as clarifying the set existence commitments of a single proof, of his Well-Ordering Theorem”.

Another element in understanding this law of thought is its connection with the significance Cantor gave to the notion of mathematical existence. For him, this last aspect implies *both* objective and subjective elements. The law of thought operates on sets, objectively existent entities. But the act of arranging is subjective.

Frege considers that we cannot always (or even very often) impose an order on extensions which are not given to us in some serial form, rejecting therefore subjective implications. He distinguishes between “subjective capacity to arrange in an order” and “existence of an order” (Frege, 1884/1980, p. 54), and shows that it is quite possible to view the existence of an order relation (and a well-ordering at that) in a completely objective way. Among the examples of “subjective contributory factors” he refers to “the amount of time at our disposal or on the extent of our familiarity with the things concerned” (Frege, 1884/1980, p. 93).

Russell adopts the same existential position (regarding number) as Frege, although not on identical lines, characterized by views that clearly dismiss our subjective capacity to order. A set is given with its orders, so the prominence of one order or another must be considered arbitrary, result of our subjective choices. It would be better therefore to consider an order as a serial relation among the terms of a set, and not a property of a set, such that, given the relation, the field (of terms) is also given, but not vive-versa: “the consideration of the terms is superfluous, and that of the relation alone is quite sufficient” (Russell, 1903/1996, p. 242).

The identity *arranging a set in an order* / *defining an order* / *proving the existence of an order* was not, as Hallett (1984) claims, characteristic of Cantor, but it was present in the way Hilbert saw the answer to the continuum problem in his 1900 lecture[8]. In other words, Hilbert maintained that the solution to the continuum hypothesis would first require a proof of the well-ordered character of the continuum.

PARADOXES

Russell’s assertion from 1903 that the set of ordinals, although linearly-ordered, is not well-ordered is rather incongruent with the fact that the initial segment of every set of ordinals is well-ordered. He eventually abided by his paradox, concentrating on, as in rejecting, the idea of the set of all sets. It was also the choice of Zermelo. The other choice, assumed by other mathematicians, is to introduce the safer concept of *class*, as I shall indicate below.

[8] For Hilbert’s lecture, see Yandell (2002), particularly pp. 386-387.

According to several accounts, it was Zermelo who first draw attention to what was later called Russell's paradox, but because he did not publish his result, his discovery remained known only to Hilbert and other members from the University of Göttingen. He first published his acknowledgment of this discovery in 1908, when he presented both another defence for the axiom of choice and the principle of well-ordering, and an axiomatization for set theory: "I had, however, discovered this antinomy myself, independently of Russell, and had communicated it prior to 1903 to Professor Hilbert among others"[9].

Zermelo maintains that the solution to Burali-Forti antinomy or the form given by Russell to the set-theoretic antinomies (Zermelo, 1908/2010, pp. 141,143) imply a preservation of the well-ordering principle together with a "restriction of the notion of set". His alternative is the axiomatization of set theory, specifically the axiom of separation, which restricts the comprehension principle to "definite"[10] properties inside a set.

Besides the correspondence between Hilbert and Frege alluding to Zermelo's discovery of the paradox before Russell, there is also a confirmation in Husserl's *Nachlass*[11].

Nonetheless, according to Felix Bernstein (2005, p. 836), Cantor was aware of the paradoxes or contradictions since 1895, although the earliest written mention we have is from a letter to Hilbert, dated 26 September 1897[12]. Apparently, he was not affected by them the same way a part of the mathematical world was affected after encountering Russell's paradox, and I believe that we could find three *fundamental* reasons for this fact. I emphasize the '*fundamental*' part because those reasons are intimately intertwined with the way Cantor conceived his theory of sets.

Case 1. The principle of comprehension:

Firstly, he did not assume the principle of comprehension[13]. In fact, in the 1885 review of *Grundlagen der Arithmetik*, he criticizes Frege's employment of this principle:

> Less successful, however, it seems to me his own attempt to strictly justify the number concept. The author has the unfortunate idea to take as basis for the

[9] Footnote 22 to "A new proof of the possibility of a well-ordering" (Zermelo, 1908/2010, p. 141).

[10] "Definite" here was to create a lot of discussions.

[11] For more on this last account, see the article of B. Rang and W. Thomas (1981), "Zermelo's discovery of the 'Russell Paradox".

[12] In a letter to Jourdain from 4 November 1903, Cantor writes that he sent his proof of the aleph theorem based on the distinction between consistent and inconsistent multiplicities to Hilbert in 1896-1897. He apparently does not precisely recollect the date, as he writes "Approximately 7 years ago …" ["Schon vor etwa 7 Jahren…]." (Meschovski & Nilson, 1991, p. 172).

[13] Loosely speaking, given a property P, there is a set of objects satisfying property P.

> number concept … that which the logic school names the «extension of a concept»; he fails to see that the «extension of a concept» is quantitatively something completely indeterminate [Unbestimmtes]; only in certain cases is the «extension of a concept» quantitatively determinate, and then, certainly, if it is finite, it has a determinate number, and if it is infinite, a determinate power. For such a quantitative determination of «extension of a concept», the concepts of «number» and «power» must previously be given, and it is a *perversion of what is right* if one attempts to ground the latter concepts on the concept of the «extension of the concept»[14]. (Cantor, 1885a/1932, p. 440)

But Ferreirós maintains that Frege's fifth law "was nothing but a principle of extensionality, while comprehension was implicit in the very notation employed by Frege", and that apparently, "nobody pinned down the crucial principle of comprehension before the emergence of the paradoxes. (Ferreirós, 2007, p. 252f.).

Case 2. Consistent and inconsistent multiplicities:

Secondly, and probably in conformity with his philosophical and theological convictions, embedded in his theory of sets, Cantor introduced a distinction between consistent and inconsistent multiplicities. Because of his Platonist approach to sets, they rather supported than threatened his view, in contradistinction to the logicist view of Dedekind, Frege, or Russell.

This decision was basically entailed by his distinctions with regard to the notion of the infinite, in particular the introduction of the concept of absolute infinite infinity (or simply the Absolute), on one hand, and his theory of cardinality, on the other hand.

2.1. The Absolute:

In the *Grundlagen*, Cantor maintains that Spinoza's rejection of infinite numbers is based on a *petition principii* in the arguments, one that goes back to Aristotle, since he is using a notion of number that tacitly assumes finiteness, and on the assumption that the absolute infinite allows no determination. Cantor's solution to the first problem is to accept certain "modifications [Modifikationen]

[14] „Weniger erfolgreich dagegen scheint mir sein eigener Versuch zu sein, den Zahlbegriff streng zu begründen. Der Verf. kommt nämlich auf den unglücklichen Gedanken… dasjenige, was in der Schullogik der «Umfang eines Begriffes» genannt wird, zur Grundlage des Zahlbegriffs zu nehmen; er übersieht ganz, daß der «Umfang eines Begriffs» quantitativ im allgemeinen etwas völlig Unbestimmtes ist; nur in gewissen Fällen ist der «Umfang eines Begriffs» quantitativ bestimmt, dann kommt ihm allerdings, wenn er endlich ist, eine bestimmte Zahl und, falls er unendlich ist, eine bestimmte Mächtigkeit zu. Für eine derartige quantitative Bestimmung des «Umfangs eines Begriffs» müssen aber die Begiffe «Zahl» und «Mächtigkeit» vorher von anderer Seite her bereits gegeben sein, und es ist eine *Verkehrung des Richtigen*, wenn man unternimmt, die letzteren Begriffe auf den Begriff «Umfang eines Begriffs» zu günden." (Cantor, 1885/1932, p. 440).

which, although not finite, are nevertheless determinable by numbers and are therefore what I call proper-infinite" (Cantor, 1883/2005, p. 891), namely the infinite numbers, the "*Transfinitum*", or the "*Suprafinitum*".

In this respect, the *Grundlagen* will develop one of the key elements of the Cantorian set theory, a theological and metaphysical argument transformed into a mathematical one (with the notion of inconsistent class, introduced in 1899), namely, the *absolute actual infinite*, a symbol for the infinite sequence of numbers, something that "can only be acknowledged [anerkannt] but never known [erkannt] – and not even approximately known." (Cantor, 1883/2005, p. 916, n. 2).

The infinity of the number classes, although graspable ideas and not representations (Vorstellungen), opens a path which can be pursued "ever further", but "we shall never reach a boundary that cannot be crossed; …we shall also never achieve even an approximate conception of the absolute" (Cantor, 1883/2005, p. 916, n. 2).

On January 22 1886, Cantor wrote to cardinal Franzelin that in addition to the difference between the infinite in *natura naturans* and in *natura naturata* (Spinoza's terms), he further distinguished between an "*Infinitum aeternum increatum sive Absolutum*", reserved for God and his attributes, and an "*Infinitum creatum sive Transfinitum*", expressed in the created nature and in the actually infinite number of objects in the universe (see Meschowski & Nilson, 1991, pp. 254-5).

For Cantor, the infinite sequence of transfinite numbers standing for the Absolute, together with the particularities regarding the inconsistent multiplicities will prove its relevancy in the face of the paradoxes: the Absolute cannot be mathematically determined, therefore the 'totality' of all ordinals or cardinals retain an indeterminate character.

In the *Grundlagen*, Cantor establishes his theory of cardinality on well-ordering. But in this paper from 1883 he does not prove that every cardinal is an aleph. In 1899, he describes it in terms of choice. In a letter from 26 September 1897 written to Hilbert, Cantor proves that every power is an aleph and differentiates between transfinite sets and absolutely infinite sets, later called (in a letter from 1899 to Dedekind) consistent, respectively, inconsistent multiplicities.

> Theorem: 'The totality of all alephs cannot be conceived as a determinate, a well-defined, and *also a finished set'*. This is the *punctum saliens*, and I venture to say that this *completely certain theorem, provable rigorously from the definition of the totality of all alephs*, is the most important and noblest theorem of set theory. One must only understand the expression 'finished' correctly (Letter to Hilbert, 2.10.1897, as cite in Ewald, 2005, pp. 927-8).

He must have been aware for a time now that no concept can be associated to an *arbitrary* set of points on the line. Hence the renunciation of any intensional approach to sets. In another letter to Hilbert he writes:

> I say of a set that it can be thought of as *finished* (and call such a set, if it contains infinitely many elements, 'transfinite' or 'super-finite') if it is possible without contradiction (as can be done with finite sets) to think of all its elements as existing together, and so to think of the set itself as a compounded thing for itself; or (in other words) if it is possible to imagine the set as actually existing with the totality of its elements.
>
> So the 'transfinite' coincides with what has since antiquity been called the 'actual infinite', and is to be considered as an ἀφωρισμένον ['something determined'].
>
> …But an 'assembling together' is only possible if an 'existing together [*Zusammensein*] is possible. (Letter to Hilbert, 2.10.1897, as cited in Ewald, 2005, pp. 927-8)

And, according to the letter to Dedekind from 3 August 1899,

> If on the other hand the totality of the elements of a multiplicity can be thought of without contradiction as 'being together', so that they can be gathered together into 'one thing', I call it a consistent multiplicity or a 'set'. (In French and in Italian this notion is aptly expressed by the words 'ensemble' and 'insieme'.) (Letter to Dedekind, 3.8.1899, as cited in Ewald, 2005, p. 932).

Another characteristic of the above-mentioned multiplicities is mentioned in another letter to Dedekind: the elements are unconnected, although they coexist:

> I wish to add that when I spoke there of multiplicities, I tacitly had in view multiplicities of *unconnected things* [Vielheiten unverbundener Dinge], that is, multiplicities such that the removal of any one or several elements has no influence on the remaining-in-existence [Bestehenbleiben] of the remaining elements (Letter to Dedekind 16.9.1899, as cited in Ewald, 2005, p. 936).

A question emerging from all this assertions is how one gets to the 'whole'. We, as thinking subjects, 'create' the unity, that is, the set, from the elements: "a many which can be thought of as a one" (Cantor, 1883/2005 and 1899). Cantor refers frequently to our capacity to create or arrange. He mentions a law of thought, which some could identify with an intension. Some interpretations consider the notion of God: since the human mind is finite, the bridge between a collection and its unity as a set is assured by God; but, because of the domain principle, Cantor's approach to sets' nature must be considered as based on extensionalism (see Hallett, 1984, pp. 34-36).

In accordance with the principle of comprehension, the paradoxes would assume the notion of a universal class, although, as the quotes above show, Cantor does not admit the universal class to be a "finished" thing or a set. But probably because of the ambiguities regarding the definition of set and Cantor's different philosophical and theological notions, the paradoxes developed in quite a problem at the turn of the century.

Cantor does not provide more explanations for the notion of set, and still today there is not a clear definition for this concept: it remained unexplained and continues, unexplainable, as a primitive. Russell, Poincaré, Brouwer, or Weyl, among others, were not too content with this unclear notion of set.

Cantor only mentions the law of thought, involving therefore the idea of an intellectual creation of unity, and left some clues in his philosophical preoccupations – his mention of Plato and his deep interests in Spinoza's philosophy.

In any case, any further elucidation must consider all the functions and roles the concept of set and set theory were supposed to assume and what they eventually accomplished: (i) the fact that set theory became the common language of mathematics (although lately other options became available, such as category theory); (ii) that it constitutes a specific branch of mathematics, with different ramifications; and (iii) that it basically offers a theory of the infinite. And the first aspect implied that different mathematical schools, each with its own philosophy and attitudes toward what mathematics is and how it should be conducted, should be provided with a concept wide enough to meet all their expectations.

This third aspect is the most philosophically charged, and it also constitutes a central reason for some conceptual assumptions, probably the most important being the fact that Cantor did not adopt an intensional approach. A set is not determined by a concept, it *is* itself a concept. The reasons might be (i) Spinoza's theory of modes and ideas; and (ii) a theologically determined notion of God, although the organic character which he ascribes to the universe of sets expressing the Absolute or to the concept of set itself also involves Spinozist influences. But another reason could be, as Ferreirós (2007, p. 73) pointed out, the fact that no concept can be associated to an arbitrary set of points in the line.

2.2. Cantor's theory of cardinality:

With regard to the idea of cardinality, Cantor does not embrace a definition in terms of equivalence classes, in the tradition of Frege and then Russell, who use the principles of comprehension and extensionality. A fundamental reason, according to Hallett (1984, p. 126), is Cantor's conception of the Absolute and the distinction between consistent and inconsistent multiplicities.

While demonstrating that the set of all alephs (ת) is an inconsistent multiplicity, Cantor emphasizes some interesting elements:

> The question now arises whether *all transfinite cardinal numbers* are contained in the system Ω[16]. In other words, is there a *set* whose power is not an aleph?
>
> This question is to be answered *negatively*, and the reason for this lies in the *inconsistency* that we discerned in the systems Ω[15] and ת.

[15] Ω represents the sequence of ordinals.

> Proof. If we take a definite multiplicity V and assume that *no aleph* corresponds to it *as its cardinal number*, we conclude that V must be *inconsistent*.
>
> For we readily see that, on the assumption made, the whole system Ω is projectible into the multiplicity V, that is, there must exist a submultiplicity V' of V that is equivalent to the system Ω.
>
> V' is *inconsistent* because Ω is, and the same must therefore be asserted of V.
>
> Accordingly, every transfinite *consistent multiplicity*, that is, every transfinite set, must have a *definite aleph* as its cardinal number. Hence
>
> C. The system ת of all alephs is nothing but the system of all transfinite cardinal numbers.
>
> All sets, and in particular all '*continua*', are therefore '*denumerable*' in an *extended sense*.
>
> Furthermore C makes it clear that I was right when I stated [1895, p. 484] the theorem:
>
> 'If $\mathfrak{a}$ and $\mathfrak{b}$ are arbitrary cardinal numbers, then $\mathfrak{a} = \mathfrak{b}$ or $\mathfrak{a} < \mathfrak{b}$ or $\mathfrak{a} > \mathfrak{b}$.'
>
> For, as we have seen, these relations of magnitude obtain between the alephs. (Letter to Dedekind, 3.08.1899, as cited in Ewald, 2005, pp. 934-5).

Some observations:

(1) Ω becomes an instrument for defining the inconsistent or absolute infinity: it is projected into the multiplicity V and equivalent to a submultiplicity V' of V.

But the ordinal numbers are for Cantor inextricably connected to well-ordering and that special law of thought. The 'description' of the Absolute must be related, in a way or another, to the way one understands well-ordering. To quote Hallett (1984, p.168), since what Cantor establishes is that "X is not a set (is an absolute collection) if and only if it has a sub-collection equivalent to the ordinal numbers sequence", we are justified to consider that "the ordinal sequence, so to speak … measure absoluteness".

In his footnote to Cantor's proof, Zermelo lays stress on the weakness of the 'projection' line of reasoning: (1) it is not certain that Ω is " 'projectible' into every multiplicity V that has no aleph as its cardinal number"; (2) Cantor uses the intuition of time in trying to establish a correspondence between the successive and arbitrary elements of V and the numbers of Ω, with every element of V used only once; (3) an exhaustive crossing of all V could lead to either a mapping on a segment of Ω and its power would be an aleph contrary to the assumption, or V would remain inexhaustible, Ω would be mapped on a subset V' of V and V would be therefore inconsistent; (4) we could define V' as a subset of V only through the axiom of choice, but Cantor uses it "unconsciously and instinctively"; (5) 'inconsistent multiplicity' represents a contradictory notion that cannot be used to prove the well-ordering theorem, unlike the axiom of choice[16] (Letter to Dedekind, 3.8.1899, as cited in Ewald, 2005, p. 935, footnote of E. Zermelo).

I do not think that Cantor can be denounced for using an intuition of time in his account of ordinal numbers, considering that Zermelo himself acknowledges the use of the well-ordering principle and hence the axiom of choice, albeit "unconsciously and instinctively". Furthermore, the theory of abstraction to which Cantor adhered in his last presentation of set theory (Cantor 1895/1897) is characterized by immediacy and simultaneity. Additionally, he explicitly rejected on many occasions any temporal aspect in his approach to the transfinite.

What countability of all continua in an extended sense means is that they can be well-ordered in a sense that involves infinite simultaneous choices, in accordance with the axiom of choice, and not the temporally determined arbitrary choices Zermelo was referring to in the footnote quoted above.

If we could adhere to the idea that the sequence Ω of all transfinite ordinals is equivalent to a submultiplicity V' of the V on account of its projection, and we also consider the principle of well-ordering, we could get to the idea of well-ordered universe (of mathematical objects)[17], but with a caveat: we never get to a whole. In other words, we reject Parmenides, but we accept Spinoza: there is one substance with an infinity of infinite modes, conceived under an infinity of attributes, but never do their totality form the entire substance.

The universe is too big to be numbered, but, because 'numbered' here cannot be treated in the traditional way, this proof could also draw some new light on the principle of well-ordering.

(2) The inconsistency of V was considered equivalent to the fact that no aleph corresponds to it as its cardinal number. That implies the fact that only (consistent) sets have alephs as cardinal numbers.

(3) A consequence of V being inconsistent is that every (all) transfinite set(s) must have an aleph as its cardinal number.

(4) As another consequence of this proof, Cantor maintains that all sets, and in particular all '*continua*', are '*denumerable*' in an *extended sense*. This, of course, implies that they actually are sets. But we still don't have a clear definition of the notion of set. In fact, in another letter to Dedekind, Cantor explicitly affirms that the consistency of both the finite and infinite sets must be postulated.

[16] Zermelo first used the axiom of choice to prove the well-ordering theorem in 1904.

[17] Hallett adjusts Cantor's arguments to a modern axiomatic theory (VNB) to emphasize the strong assumptions of Cantor in this proof (see Hallett 1984, pp. 172-4), since "The VBN reconstruction shows that (local) well-ordering together with this structural assumption [of the role played by the ordinals] lead directly to the well-orderability of the universe" (Hallett, 1984, p. 173).

Set must therefore be taken as primitive notion. Hence the axiomatizations. And eventually, to every axiomatization we apply Gödel's results and accept that no axiomatized system exhausts our mathematical knowledge, that we cannot know everything there is to know about numbers, sets, or any other mathematical object, the relations among them or the consequences of these relations.

Cantor used therefore his result that the ordinal numbers form an inconsistent multiplicity to show that the cardinal numbers are well-ordered, statement identical to the fact that every cardinal number is an ℵ, the power of the set of predecessors of an ordinal number. He took this result as an axiomatic manifestation of a "transfinite arithmetic", an extension of the old arithmetic18, in which every natural number is a finite cardinal: another anticipation of some future axioms in the modern set theory. If the cardinality of a multiplicity is not an ℵ, the entire sequence of ordinals would be projectible into that multiplicity and the nature of that multiplicity would be unveiled as inconsistent, therefore not as a set.

In other words, Cantor based his theory of power or cardinality upon the principle of well-ordering, such that the order of the cardinal numbers uses the inherent order of the ordinal numbers. That is why we say today that well-ordered cardinals are identified with their initial ordinals, i.e., the smallest ordinal of that cardinality (all infinite cardinals are limit ordinals).

In 1915, Hartogs proved that any system of cardinally comparable sets is based on the well-ordering theorem (proved by Zermelo in 1904 and 1908), relation defined as equivalence. In Zermelo-Fraenkel system of axioms, the cardinal comparability could be based on the von Neumann cumulative hierarchy.

Today, that is in ZFC (the Zermelo-Fraenkel system of axioms together with the axiom of choice), the aleph function is considered a bijection between transfinite ordinals and cardinals. Any infinite ordinal has an aleph as cardinality. That implies that every set whose cardinality is an aleph is equipollent with an ordinal and therefore well-ordered. A finite set is also well-ordered, but it doesn't have an aleph as cardinality. On the other hand, in ZF without the axiom of choice, we cannot prove that every set has an aleph as cardinality.

Case 3. The power set:

Thirdly, the apparent paradox draws attention to a central aspect of set theory: the existence of power sets.

[18] "In other words: the fact of the 'consistency' of finite multiplicities is a simple, unprovable truth; it is 'The axiom of arithmetic (in the old sense of the word)'. And in the same way the 'consistency' of the multiplicities to which I assign the alephs as cardinal numbers is 'the axiom of extended, of transfinite arithmetic'" (Letter to Dedekind, 28.8.1899, as cited in Ewald, 2005, p. 937).

On the whole, not only because of the Platonist view with regard to the nature of sets did Cantor adopt a detached ‘acceptance’ of the paradoxes. I believe that another element was his construction of the power set (Cantor, 1891/2005) – the set of all subsets of a given set – which offered a new way, somehow disruptive, of conceiving sets. Cantor did not actually used the term ‘power set’: the ‘power’ from what we now call power-set translates the German word “Potenz”, employed for exponentiation.

The power set was used in the 1891[19] article to prove, through the diagonal argument, what we now call the Cantor’s Theorem or the fact that the power set of a given set L has a greater cardinality than the set L itself[20]. Russell got his paradox by applying the theorem to the set of all sets.

A. Kanamori (2009, p. 402) maintains that “it would be an exaggeration to assert that Cantor was working on power sets; rather, he had expanded the 19th century concept of function by ushering in arbitrary functions”. Nevertheless, although he did not employ the modern term, it can be considered that he did capture the concept, as some of his letters underline. In them, Cantor uses the notion of ‘completed’ or ‘finished’[21] (*fertig*) set: “«If M is a finished set, so is each subset of M a completed set» … «The collection of all subsets of a completed set M is a completed set». For *all the subsets* of M are «together» contained in M; the fact that they might coincide, does no harm.”[22] (Letter to Hilbert, 10.10.1898, as cited in Meschkowski & Nilson, 1991, p. 396).

In the same letter, after introducing the set S of all the functions $f(v)$ with values 0 or 1 (the characteristic function), Cantor states the equivalence[23] with the linear continuum (*Linearcontinuum*) and considers it a “completed set” (*fertige Menge*)[24]. In this letter he explicitly considers the linear continuum as a completed set[25], but in a letter from 9 May the following year, he expresses his doubts (Letter to Hilbert, 9 May 1899, as cited in Meschkowski & Nilson, 1991, p. 399). These last aspects emphasize the fact that Cantor’s fear might not have been the

[19] “Über eine elementare Frage der Mannigfaltigkeitslehre”.

[20] In the 1895 notation, $\aleph < 2^{\aleph_0}$.

[21] Ewald translates it with “finished” (see Ewald, 2005, p. 927).

[22] “Ist M eine fertige Menge, so ist auch jede Theilmenge von M eine fertige Menge. … «Die Vielheit *aller Theilmengen* einer fertigen Menge M is einer fertige Menge». Denn alle Theilmengen von M sind «zusammen»in M enthalten; der Umstand, daß sie sich theiweise decken, schadet hieran nichts.“ (Meschkowski & Nilson, 1991, p. 396).

[23] A bijective correspondence.

[24] „Das Linearcont ist aequivalent der Menge S = {f(v)} wo f(v) die Werthe 0 oder 1 haben kann. … Ich behaupte also S ist eine «fertige Menge»“ (Meschkowski & Nilson, 1991, p. 396).

[25] Which is a completed set: „Das Linearcontinuum is eine fertige Menge” (Meschkowski & Nilson, 1991, p. 396).

paradoxes, but the fact that the continuum reformulated as a power-set is not well-ordered.

The introduction of what we now call the power set is fundamental from several points of view. Firstly, it is tightly connected to the idea of mathematical existence, in particular, the existence of real numbers. As S. Lavine (1994, p. 95) rightfully points out, it proves "*for the first time that the real numbers form a set*, instead of just taking that as an additional assumption" [*emphasis added*], providing "a new independent argument for the existence of the real and transfinite numbers".

Secondly, and also subscribing to the observations of the author mentioned above, the concept of power set determined a breach in the, until then, Cantor's unified notion of set, countable because well-ordered: by the construction of power sets "as a second set-existence principle, Cantor no longer had a unitary conception of set. He could therefore no longer say which multiplicities were sets; the only way he had to show that some were not sets was by arriving at contradictions" (Lavine, 1994, p. 98). The well-ordering of the continuum was not no longer definable.

We well-order the set of all countable ordinals, $\aleph 1$, such that it comes after the cardinality of each ordinal from the second number-class, but it is not so obvious where in this order is $2^{\aleph_0}$ situated, on account of the fact that they are created by two different processes. Hence the continuum hypothesis.

The creation of the power set is also a thought process, as Cantor himself maintains, reflecting the ability of our mind to conceive such objects. This process of creation can be continued *ad infinitum*, with the result of an infinite hierarchy of infinite sets, subject to a specific arithmetic, known as the transfinite arithmetic. The question of their *existence* opens other problems.

FINAL REMARKS

It must also be emphasized that an inconsistent multiplicity is incompatible with the definition of set (Menge) from the *Beiträge* as "any collection into a whole (*Zusammenfassung zu einem Ganzen*) M of definite and separate objects *m* of our intuition or our thought." (Cantor, 1885/1887, p. 85, § 1). This paper is generally considered Cantor's last word on his theory of sets, disregarding therefore the later correspondence with different mathematicians.

But even considering – although disregarding – the letters, Cantor's set theory is also commonly – and wrongfully – regarded as naïve. Purkert & Ilgauds (1987, pp. 150-9) or Ferreirós (2007) are some of those who reject this categorization of Cantor's set theory. As Ferreirós maintains, "[T]he naive standpoint can be found explicitly in Frege and Russell, implicitly in Riemann and Dedekind[26]. But the frequent confusion is understandable, for Cantor's subtle distinction had not been sufficiently clarified at all." (Ferreirós, 2007, p. 292).

The paradoxes do affect the notion of cardinality assumed by Frege and Russell as the set of all sets equinumerous with M, because of the 'the set of all sets' assumption. Not every property P determines a consistent set, as the principle of comprehension requires, following the traditional logic, where the leap from a property to a class or multitude was somehow natural.

A modern alternative in circumventing multitudes too big to be considered inconsistent is the distinction (attributed to von Neumann) between sets and classes. We could consider every set to be a class. A class which is not a set is a proper class. Informally, any collection of the form $\{x; \varphi(x)\}$ is a class, a proper class being a class that does not form a set because it is "too big" (Kunen, 1980, p. 23). According to the Axiom Schema of Comprehension or Specification, any subclass of a set is a set.

Formally, proper classes do not exist, and an expression involving them must be thought of as abbreviations for expressions not involving them. Thus $x \in ON$ abbreviates the formula expressing that x is an ordinal, and $ON = V$ abbreviates the (false) sentence (abbreviated by)

$\forall x$ (x is an ordinal $\leftrightarrow x=x$). … Any of our defined predicates and functions might be thought of as a class. (Kunen, 1992, p. 24).

As a consequence, the paradoxes which apparently frightened a great number of mathematicians with logicist tendencies were more likely to offer new perspectives on some fundamental notions of Cantor's set theory, to highlight their range of applicability, particularly regarding the concept of set. Cantor offered the distinction between consistent and inconsistent multiplicities in 1899, a few years after he discovered the paradoxes, or after they were publically mentioned. The third principle of generation, which imposes certain constraints on the endless process of transfinite creation, was already introduced in the *Grundlagen*, in 1883.

The introduction of the power set did dot annul this earlier form of the limitation of size principle[27], but probably offered a new condition that strengthened it. Both Hallett (1984) and Ferreirós (2007) maintain that Cantor's analysis from the 1899 letters (and we can also include that from 1897 to Hilbert) anticipate some of the Zermelo-Fraenkel axioms: the Separation axiom (every part of a set is a set), the Replacement axiom (two equipollent classes are either both sets, or both inconsistent), the Union axiom (the union of a sets is a set)[28].

[26] In fact, in the letter from 1899, Cantor shows that the paradoxes affect Dedekind's logicist sets (Dedekind used the term "system"), and the conception of the totality of everything thinkable (another notion belonging to Dedekind, see Ferreiros, 2007, p. 292).

[27] The 'limitation of size' was a notion introduced by Russell in 1906 as a possible solution to the paradoxes.

[28] See (Ferreirós 2007, p. 294).

REFERENCES

Bernstein, F. (2005). Account. In W. Ewald (Ed.), *From Kant to Hilbert: A Source Book in the Foundations of Mathematics* (p. 836). II. Oxford: Clarendon Press.

Burali-Forti, C. (1967a). A Question on Transfinite Numbers. In J. van Heijenoort (Ed.), *From Frege to Gödel. A Source Book in Mathematical Logic.1879-1931* (pp. 105-111). Cambridge: Harvard University Press (Original work published in March, 1897).

Burali-Forti, C. (1967b). On well-ordered classes. In J. van Heijenoort (Ed.), *From Frege to Gödel. A Source Book in Mathematical Logic. 1879-1931* (pp. 111-112). Cambridge: Harvard University Press. (Original work published in October, 1897).

Cantor, G. (1932). Über unendliche, lineare Punktmannigfaltigkeiten, 3. In E. Zermelo (Ed.), *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts* (pp. 149-57). Berlin: Springer. (Original work published in 1882).

Cantor, G. (2005). Grundlagen einer allgemeinen Mannigfaltigkeitslehre. Ein mathematisch philosophischer Versuch in der Lehre des Unendlichen. In W. Ewald (Ed. and Trans.), *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*. II (pp. 878-920). Oxford: Clarendon Press (Original work published in 1883).

Cantor, G. (1932). "Rezension der Schrift von G. Frege Die Grundlagen der Arithmetic", *Deutsche Literaturzeitung*, 6. Jahrgang, pp. 728-9. In E. Zermelo (Ed.), *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts* (pp. 440-1). Berlin: Springer (Originally published in 1885).

Cantor, G. (1932). Mitteilungen zur Lehre vom Transfiniten II, *Zeitschrift fur Philosophie und philosophische Kritik* 91, 81-125, 252-70; 92, 250-65. In E. Zermelo (Ed.), *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts* (pp. 378-439). Berlin: Springer (Originally published in 1887-1888).

Cantor, G. (2005). 1891: Über eine elementare Frage der Mannigfaltigkeitslehre, *Jahresbericht der deutschen Mathematiker-Vereiningung* 1, 75-8. In W. Ewald (Ed. and Trans.), *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*. II (pp. 920-922). Oxford: Clarendon Press (Originally published in 1891).

Cantor, G. (2010). Beiträge zur Begründung der transfiniten Mengenlehre, 1&2. In *Contributions to the Founding of the Theory of Transfinite Numbers* (pp. 85-136). New York: Dover Publications (Originally published in 1895 & 1897).

Ewald, W. (Ed.) (2005). *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*. II. Oxford: Clarendon Press.

Ferreirós, J. (2007). *Labyrinth of Thought. A History of Set Theory and Its Role in Modern Mathematics*. Basel: Birkhäuser.

Frege, G. (1964). *The Basic Laws of Arithmetic. Exposition of the System* (M. Furth, translator). Berkeley: University of California Press (Originally published in 1893).

Frege, G. (1980). *The Foundations of Arithmetic. A logico-mathematical enquiry into the concept of number* (J. L. Austin, translator). New York: Harper Brothers (Originally published in 1884).

Hallett, M. (1984). *Cantorian set theory and limitation of size*. Oxford: Clarendon Press.

Heijenoort, J. van (1967). *From Frege to Gödel. A Source Book in Mathematical Logic.1879-1931*. Cambridge: Harvard University Press.

Kanamori, A. (2009). Set theory from Cantor to Cohen. In A. D. Irvine (Ed.), *Philosophy of Mathematics* (pp. 395-460). Amsterdam: North Holland.

Kunen, K. (1992). *Set Theory. An Introduction to Independence Proofs*. Amsterdam: North Holland.

Lavine, S. (1994). *Understanding the Infinite*. Cambridge, Massachusetts: Harvard University Press.

Meschkowski, H., & Nilson, W. (Eds.). (1991). *Georg Cantor: Briefe*. Berlin: Springer.

Purkert, W., & Ilgauds, H. J. (1985). *Georg Cantor: 1845-1918*. Basel: Birkhäuser Verlag.

Rang, B., & Thomas, W. (1981). Zermelo's discovery of the 'Russell Paradox'. *Historia Mathematica*, *8*(1), 15–22.

Russell, B. (2010). *Introduction to mathematical philosophy*. London: George Allen & Unwin, Ltd. Online Corrected Edition version 1.0. (Originally published in 1919)

Russell, B.(1996). *The Principles of Mathematics*. New York: W. W. Norton & Comp. (Originally published in 1903).

Spinoza, B. de (2002). *Spinoza. Complete Works* (S. Shirley, translator). Indianapolis/Cambridge: Hackett Publishing Company.

Yandell, B. H. (2002). *The Honors Class. Hilbert's Problems and Their Solvers*. Natick, Massachusetts: A K Peters.

Zermelo, E. (2010). *Collected Works. Gesammelte Werke*. Berlin Heidelberg: Springer-Verlag (Originally published in 1908).